

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number
WO 01/56216 A2

(51) International Patent Classification⁷: H04L

(21) International Application Number: PCT/US01/02316

(22) International Filing Date: 24 January 2001 (24.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/178,077 25 January 2000 (25.01.2000) US

(71) Applicant (for all designated States except US):
AFFYMETRIX, INC. [US/US]; 3380 Central Expressway, Santa Clara, CA 95051 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): CRAFORD, David, M. [US/US]; 1027 Ray Avenue, Los Altos, CA 94022 (US).
NORVIEL, Vernon, A. [US/US]; 1155 El Abra Way, San Jose, CA 95125 (US).

(74) Agent: SHERR, Alan, B.; Affymetrix, Inc., 34 Commerce Way, Woburn, MA 01801 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

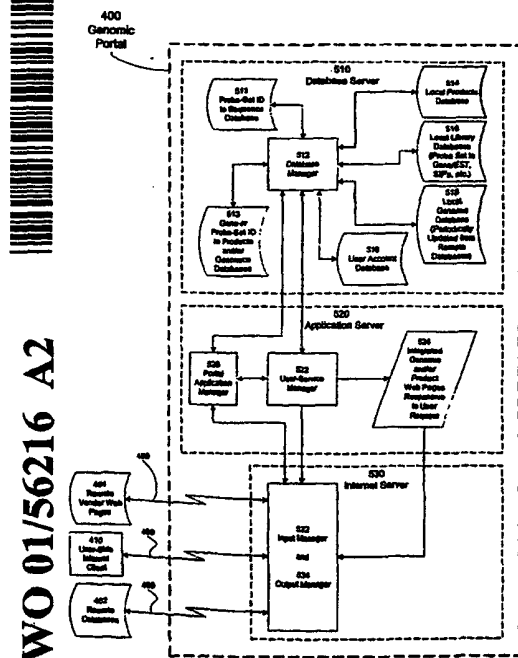
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD, SYSTEM AND COMPUTER SOFTWARE FOR PROVIDING A GENOMIC WEB PORTAL



(57) Abstract: Systems, methods, and computer program products are described that process inquiries or orders regarding purchase of biological devices, substances, or related reagents. In some implementations, a user selects probe-set identifiers that identify microarray probe sets capable of enabling detection of biological molecules. Corresponding genes or EST's are identified and are correlated with related product data, which is provided to the user. Further, the user may select products for purchase based on the product data. If so, the user's account may be adjusted based on the purchase order. In the same or other implementations, a local genomic database is periodically updated. In response to a user selection of probe-set identifiers, data related to corresponding genes or EST's is provided to the user from the local genomic database.

WO 01/56216 A2

METHOD, SYSTEM, AND COMPUTER SOFTWARE FOR PROVIDING A
GENOMIC WEB PORTAL

5 RELATED APPLICATION

The present application claims priority from U.S. Provisional Patent Application Serial No. 60/178,077, entitled "METHOD, SYSTEM, AND COMPUTER SOFTWARE FOR PROVIDING A GENOMIC WEB PORTAL," filed January 25, 2000,
10 incorporated herein by reference in its entirety for all purposes.

BACKGROUND

The present invention relates to the field of
15 bioinformatics. In particular, the present invention relates to computer systems, methods, and products for providing genomic information over networks such as the Internet.

Research in molecular biology, biochemistry, and
20 many related health fields increasingly requires organization and analysis of complex data generated by new experimental techniques. These tasks are addressed by the rapidly evolving field of bioinformatics. See, e.g., H. Rashidi and K. Buehler, Bioinformatics Basics: Applications in Biological Science and Medicine (CRC Press, London, 2000); Bioinformatics: A Practical Guide to the Analysis of Gene and Proteins (B.F. Ouellette and A.D. Bzevanis, eds., Wiley & Sons, Inc., 1998), both of which are hereby incorporated herein by reference in
25 their entireties. Broadly, one area of bioinformatics applies computational techniques to large genomic databases, often distributed over and accessed through
30 networks such as the Internet, for the purpose of

- 3 -

researchers with improved tools and information to perform these tasks.

Systems, methods, and computer program products are described herein to address these and other needs. In some implementations, a web portal processes inquiries or orders regarding purchase of biological devices or substances, or related reagents. The user selects "probe-set identifiers" (a broad term that is described below) that may be associated with probe sets of one or more probes. These probe sets are capable of enabling detection of biological molecules. These biological molecules include, but are not limited to, nucleic acids including DNA representations or mRNA transcripts and/or representations of corresponding genes (such nucleic acids are hereafter, for convenience, referred to simply as "mRNA transcripts"). The corresponding genes or EST's are identified and are correlated with related data, which is provided to the user. In some aspects, the user may select products for purchase based on the data. If the user decides to make a purchase, the user's account may be adjusted based on the purchase order.

An advantage of some of these implementations is that a user may be presented with product suggestions for follow-up experiments based on results from an initial experiment. These initial results are represented by the user's selection of probe-set identifiers by, for example, designating those probe-set identifiers corresponding to probes indicating a relatively high degree of differential expression in control and experimental samples.

In the same or other implementations, a local genomic database is periodically updated. In some aspects, this updating may be made from remote databases.

In response to a user selection of probe-set identifiers, data related to genes or EST's are provided

In some implementations a system is described for providing data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and
5 capable of enabling detection of a biological molecule. The biological molecule may be a nucleic acid or an mRNA transcript of a corresponding gene. As noted above, one or more of the probe-set identifiers may include a gene or EST identifier, such as an accession number. The
10 system includes an input manager that receives a user selection of a first set of probe-set identifiers; a gene determiner that identifies genes or EST's corresponding to the probe sets identified by the first set of probe-set identifiers; a correlator that correlates the genes
15 or EST's with data; and an output manager that provides the data to the user. The input and output managers of these implementations may be coupled to the user via the Internet.

The first set of probe-set identifiers may be a
20 subset of a second set of probe-set identifiers of probe sets that have enabled detection of the expression or differential expression of their corresponding genes or EST's. For example, the user may have selected the subset using a graphical user interface provided by a
25 probe-array software application. This selection may be made, for instance, by drawing a loop around out-liers in a scatter plot representation of probe sets, where the out-liers indicate probe sets having a relatively high degree of differential expression. As another of many
30 possible examples, the user may select the subset by highlighting entries of probe-set identifiers in an ordered table.

The probe sets typically are disposed on one or more probe arrays that, as noted, may be any of various types
35 of microarrays such as those synthesized using VLSIPS™

- 7 -

the steps of: receiving at a user computer a user selection of a first set of one or more probe-set identifiers, wherein each probe-set identifier identifies a probe set that has enabled detection of the expression of a corresponding gene; providing the user selection over the Internet to a portal system capable of correlating data with one or more genes or EST's corresponding to the probe sets identified by the first set of probe-set identifiers; and receiving the correlated data from the portal system. The user may also select product data for purchase.

Yet another implementation is directed to a system for providing data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of a biological molecule. The biological molecule may be a nucleic acid or an mRNA transcript of a corresponding gene. The system includes a database manager that periodically updates a local genomic database comprising data related to the genes or EST's; an input manager that receives a user selection of probe-set identifiers; a user-service manager that constructs from the local genomic database data related to genes or EST's corresponding to the probe-set identifiers; and an output manager that provides the data to the user.

In the preceding implementations, the database manager may periodically update the local genomic database, for example, weekly, with sequence data, exonic structure or location data, splice-variants data, marker structure or location data, polymorphism data, homology data, protein-family classification data, pathway data, alternative-gene naming data, literature-recitation data, annotation data, other genomic or proteomic data, or any combination thereof. This updating may be accomplished

- 9 -

providing product data, e.g., biological product data. The method includes the steps of: receiving from a user a gene, EST, and/or probe-set identifier; correlating the gene, EST, and/or probe-set identifier with one or more
5 product data; causing the product data to be obtained from a local and/or a remote database or other local and/or remote source of data; and providing the product data to the user. The method may optionally include adjusting a user account based on the purchase, or
10 adjusting a vendor account for referring the user to the vendor.

A further aspect is a system for providing product data related to one or more genes or EST's. Each gene or EST has at least one corresponding probe set identified
15 by a probe-set identifier and capable of enabling detection of a biological molecule. The system includes an input manager that receives one or more of the probe-set identifiers; a correlator that correlates the probe-set identifiers with a first set of one or more product
20 data; and an output manager that provides the first set of data to the user. Yet another aspect is a system for providing product data related to one or more genes or EST's. The system includes an input manager that receives one or more gene and/or EST identifiers; a
25 correlator that correlates the identifiers with a first set of one or more product data; and an output manager that provides the first set of data to the user.

An additional aspect is a method for providing product data related to one or more genes or EST's. Each
30 gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of a biological molecule. The method includes the steps of receiving one or more of the probe-set identifiers; correlating the probe-set identifiers
35 with a first set of one or more product data; and

- 11 -

from a user a selection of a first set of one or more of the probe-set identifiers; data managing means for periodically updating from the local genomic database a first set of data related to genes or EST's corresponding
5 to the first set of probe-set identifiers; and providing means for providing the first set of data to the user.

The above implementations are not necessarily inclusive or exclusive of each other and may be combined in any manner that is non-conflicting and otherwise
10 possible, whether they be presented in association with a same, or a different, aspect or implementation. The description of one implementation is not intended to be limiting with respect to other implementations. Also, any one or more function, step, operation, or technique
15 described elsewhere in this specification may, in alternative implementations, be combined with any one or more function, step, operation, or technique described in the summary. Thus, the above implementations are illustrative rather than limiting.

20

BRIEF DESCRIPTION OF THE DRAWINGS

The above and further advantages will be more clearly appreciated from the following detailed description when taken in conjunction with the
25 accompanying drawings. In the drawings, like reference numerals indicate like structures or method steps and the leftmost one or two digits of a reference numeral indicates the number of the figure in which the referenced element first appears (for example, the
30 element 180 appears first in Figure 1 and element 1020 first appears in Figure 10). In functional block diagrams, rectangles generally indicate functional elements, parallelograms generally indicate data, rectangles with curved sides generally indicate stored
35 data, rectangles with a pair of double borders generally

- 13 -

Figure 7 is a flow chart of one embodiment of a method for providing a user with genomic product information related to gene expression, or differential expression, experimental results;

5 Figure 8 is a functional block diagram of one embodiment of a user-service manager application as may be executed on the portal application computer system of Figure 5;

10 Figure 9 is a simplified graphical representation of one embodiment of a gene or probe-set identifier to database such as may be by the user-service manager of Figure 8 in connection with the method of Figure 7;

15 Figure 10 is one embodiment of a graphical user interface that may be generated by a probe-array analysis application of Figure 2; and

Figure 11 is another embodiment of a graphical user interface that may be generated by a probe-array analysis application of Figure 2.

20

DETAILED DESCRIPTION

Systems, methods, and computer products are now described with reference to an illustrative embodiment referred to as genomic portal 400. Portal 400 is shown in an Internet environment in Figure 4, and is
25 illustrated in greater detail in Figures 5-11.

In a typical implementation, portal 400 may be used to provide a user with information related to results from experiments with probe arrays. The experiments often involve the use of scanning equipment to detect
30 hybridization of probe-target pairs, and the analysis of detected hybridization by various software applications, as now described in relation to Figures 1 and 2.

Probe Arrays 103

Various techniques and technologies may be used for
35 depositing or synthesizing dense arrays of biological

probe array. Some targets hybridize with at least partially complementary probes and remain at the probe locations, while non-hybridized targets are washed away.

These hybridized targets, with their "tags" or "labels,"
5 are thus spatially associated with the targets' complementary probes. The hybridized probe and target may sometimes be referred to as a "probe-target pair." Detection of these pairs can serve a variety of purposes, such as to determine whether a target nucleic acid has a
10 nucleotide sequence identical to or different from a specific reference sequence. See, for example, U.S. Patent No. 5,837,832, referred to and incorporated above.

Other uses include gene expression monitoring and evaluation (see, e.g., U.S. Patent No. 5,800,992 to
15 Fodor, et al.; U.S. Patent No. 6,040,138 to Lockhart, et al.; and International App. No. PCT/US98/15151, published as WO99/05323, to Balaban, et al.), genotyping (U.S. Patent No. 5,856,092 to Dale, et al.), or other detection of nucleic acids. The '992, '138, and '092 patents, and
20 publication WO99/05323, are incorporated by reference herein in their entirety for all purposes.

Other techniques exist for depositing probes on a substrate or support. For example, "spotted arrays" are commercially fabricated on microscope slides. These
25 arrays consist of liquid spots containing biological material of potentially varying compositions and concentrations. For instance, a spot in the array may include a few strands of short oligonucleotides in a water solution, or it may include a high concentration of
30 long strands of complex proteins. The Affymetrix® 417™ Arrayer is a device that deposits a densely packed array of biological material on a microscope slide in accordance with these techniques, aspects of which are described in PCT Application No. PCT/US99/00730
35 (International Publication Number WO 99/36760), hereby

- 17 -

"probe" is not limited to probes immobilized in array format. Rather, the functions and methods described are also useful for providing genomic information and intelligent e-commerce for other parallel assay devices.

5 For example, these functions and methods may be applied with respect to probe-set identifiers that identify probes immobilized on or in beads, optical fibers, or other substrates or media.

Probes typically are able to detect the expression
10 of corresponding genes or EST's by detecting the presence or abundance of mRNA transcripts present in the target. This detection may, in turn, be accomplished by detecting labeled cRNA that is derived from cDNA derived from the mRNA in the target. In general, a probe set contains
15 sub-sequences in unique regions of the transcripts and does not correspond to a full gene sequence. The word "set" generally is used herein to refer to one or more; e.g., a probe set may consist of one or more probes, and a set of probe-set identifiers may consist of one or more
20 probe-set identifiers.

Scanner 190

Figure 1 is a functional block diagram of a system that is suitable for, among other things, analyzing probe
25 arrays that have been hybridized with labeled targets. Representative hybridized probe arrays 103 of Figure 1 may include probe arrays of any type, as noted above. Labeled targets in hybridized probe arrays 103 may be detected using various commercial devices, referred to
30 for convenience hereafter as "scanners." An illustrative device is shown in Figure 1 as scanner 190. Scanners image the targets by detecting fluorescent or other emissions from the labels, or by detecting transmitted, reflected, or scattered radiation. These processes are
35 generally and collectively referred to hereafter for

- 19 -

emissions were scanned. The pixel may also have another value representing another characteristic, such as color.

For instance, a scanned elemental sub-area in which high-intensity emissions were detected may be represented by a pixel having high luminance (hereafter, a "bright" pixel), and low-intensity emissions may be represented by a pixel of low luminance (a "dim" pixel). Alternatively, the chromatic value of a pixel may be made to represent the intensity, color, or other characteristic of the detected emissions. Thus, an area of high-intensity emission may be displayed as a red pixel and an area of low-intensity emission as a blue pixel. As another example, detected emissions of one wavelength at a particular sub-area of the substrate may be represented as a red pixel, and emissions of a second wavelength detected at another sub-area may be represented by an adjacent blue pixel. Many other display schemes are known.

Probe-Array Analysis Applications 199

Generally, a human being may inspect a printed or displayed image constructed from the data in an image file and may identify those cells that are bright or dim, or are otherwise identified by a pixel characteristic (such as color). However, it frequently is desirable to provide this information in an automated, quantifiable, and repeatable way that is compatible with various image processing and/or analysis techniques. For example, the information may be provided for processing by a computer application that associates the locations where hybridized targets were detected with known locations where probes of known identities were synthesized or deposited. Information such as the nucleotide or monomer sequence of target DNA or RNA may then be deduced.

Techniques for making these deductions are described, for example, in U.S. Patent No. 5,733,729 to Lipshutz, which

- 21 -

or EST identifiers. All of the patent applications noted in this paragraph are hereby incorporated herein by reference in their entireties.

For convenience of reference, these types of computer software applications (i.e., for acquiring and processing image files, data mining, data management, and various database and other applications related to probe-array analysis) are generally and collectively represented in Figure 1 as probe-array analysis applications 199. Figure 2 is a functional block diagram of probe-array analysis applications 199 as illustratively stored for execution (as executable code 199A corresponding to applications 199) in system memory 120 of user computer 100 of Figure 1.

As will be appreciated by those skilled in the relevant art, it is not necessary that applications 199 be stored on and/or executed from computer 100; rather, some or all of applications 199 may be stored on and/or executed from an applications server or other computer platform to which computer 100 is connected in a network.

For example, it may be particularly advantageous for applications involving the manipulation of large databases, such as Affymetrix® LIMS or Affymetrix® Data Mining Tool (DMT), to be executed from a database server such as user database server 412 of Figure 4.

Alternatively, LIMS, DMT, and/or other applications may be executed from computer 100, but some or all of the databases upon which those applications operate may be stored for common access on server 412 (perhaps together with a database management program, such as the Oracle® 8.0.5 database management system from Oracle Corporation). Such networked arrangements may be implemented in accordance with known techniques using commercially available hardware and software, such as those available for implementing a local-area network or

- 23 -

internal library files (not shown) that specify details regarding the sequences and locations of probes and controls. The goals of programs such as data analysis program 210 of this example is generally to provide
5 information such as the degree of hybridization, absolute and/or differential (over two or more experiments) expression, genotype comparisons, detection of polymorphisms and mutations, and other analytical results. In this example, file 215 represents this
10 analytical output of data analysis program 210. Data analysis program 210 may process file 215 to create report files 214 that may be responsive to requests by user 101 regarding form and content. As will be appreciated by those skilled in the relevant art, the
15 preceding and following descriptions of files, reports, and data representations generated by illustrative data analysis program 210 are exemplary only, and the data described, and other data, may be processed, combined, arranged, and/or presented in many other ways.

20 Data analysis program 210 also generates various types of plots, graphs, tables, and other tabular and/or graphical representations of analytical data such as contained in file 215. An illustrative example is shown in Figure 10, which shows a graphical user interface
25 (GUI) 1000 having scatter plot window 1010 and tabular window 1020. In scatter plot window 1010, lines 1011 provide a reference to the degree of differential expression as measured by probe sets in different experiments. The location of dots, each representing a
30 probe set from one or more microarrays, specifies along one axis the degree of expression of the probe set in one experiment or set of experiments (for example, experiments measuring control samples) and, along the other axis, the degree of expression in another

- 25 -

in some cases, indicates the accession number of the gene or EST corresponding to the probe set. For example, the probe-set identification name "M13903_at" in row 1021 indicates that the accession number of the gene or EST corresponding to the probe set corresponding to that row is M13903. In other examples, the corresponding accession number may be displayed directly. The provision of these accession numbers for selection by user 101 is represented by accession numbers 124 in Figure 2. Although, as noted, accession numbers may serve as a type of probe-set identifier (and thus accession numbers 124 may be considered as a subset of probe-set identifiers 222), they are shown distinctly in Figure 2 for convenience of illustration and discussion.

Other of applications executables 199A, such as data mining tool 220, may also provide probe-set identifiers 222 (optionally including accession numbers 224) to user 101. A further example is database application 230, an illustrative GUI of which is represented in Figure 11.

Database application 230 is an application for associating probe sets, typically identified by probe-set identifiers such as names, numbers, and/or symbols, with corresponding genes or EST's. One example of database 230 is the EASI database application from Affymetrix, noted above. In the example of Figure 11, GUI 1100 includes a query window 1110 and a results window 1120. As shown in Figure 11, user 101 has effectively created a query, in accordance with known techniques, by selecting a particular probe array 1112 and a portion 1114 of a descriptive text associated with array 1112 or any probe set associated with array 1112. Application 230 conducts a search of its database (not shown) and displays the results of the query in window 1120. As noted below with respect to database Figure 5, the functions of database application 230 and its associated database may also, or

- 27 -

by database manager 512, which may include a LIMS such as LIMS 225 or incorporate some or all of its functions.

User Computer 100

User computer 100, shown in Figure 1, may be a
5 computing device specially designed and configured to support and execute some or all of the functions of probe array applications 199. Computer 100 also may be any of a variety of types of general-purpose computers such as a personal computer, network server, workstation, or other
10 computer platform now or later developed. Computer 100 typically includes known components such as a processor 105, an operating system 110, a graphical user interface (GUI) controller 115, a system memory 120, memory storage devices 125, and input-output controllers 130. It will
15 be understood by those skilled in the relevant art that there are many possible configurations of the components of computer 100 and that some components that may typically be included in computer 100 are not shown, such as cache memory, a data backup unit, and many other
20 devices. Processor 105 may be a commercially available processor such as a Pentium® processor made by Intel Corporation, a SPARC® processor made by Sun Microsystems, or it may be one of other processors that are or will become available. Processor 105 executes operating
25 system 110, which may be, for example, a Windows®-type operating system (such as Windows NT® 4.0 with SP6a) from the Microsoft Corporation; a Unix® or Linux-type operating system available from many vendors; another or a future operating system; or some combination thereof.
30 Operating system 110 interfaces with firmware and hardware in a well-known manner, and facilitates processor 105 in coordinating and executing the functions of various computer programs that may be written in a variety of programming languages. Operating system 110,
35 typically in cooperation with processor 105, coordinates

- 29 -

Implementation of the hardware state machine so as to perform the functions described herein will be apparent to those skilled in the relevant arts.

Input-output controllers 130 could include any of a variety of known devices for accepting and processing information from a user, whether a human or a machine, whether local or remote. Such devices include, for example, modem cards, network interface cards, sound cards, or other types of controllers for any of a variety of known input devices 102. Output controllers of input-output controllers 130 could include controllers for any of a variety of known display devices 180 for presenting information to a user, whether a human or a machine, whether local or remote. If one of display devices 180 provides visual information, this information typically may be logically and/or physically organized as an array of picture elements, sometimes referred to as pixels. Graphical user interface (GUI) controller 115 may comprise any of a variety of known or future software programs for providing graphical input and output interfaces between computer 100 and user 101, and for processing user inputs. In the illustrated embodiment, the functional elements of computer 100 communicate with each other via system bus 104. Some of these communications may be accomplished in alternative embodiments using network or other types of remote communications.

As will be evident to those skilled in the relevant art, applications 199, if implemented in software, may be loaded into system memory 120 and/or memory storage device 125 through one of input devices 102. All or portions of applications 199 may also reside in a read-only memory or similar device of memory storage device 125, such devices not requiring that applications 199 first be loaded through input devices 102. It will be

- 31 -

User 101 typically copies sequence information from sequences 316 and pastes this information into an HTML document accessible through NCBI's BLAST web pages 324 (as of January 2001, accessible at
5 <http://www.ncbi.nlm.nih.gov/BLAST/>). This operation, which also may be time consuming and tedious if many sequences are involved, is represented by user-initiated batch BLAST request 322 of Figure 3. BLAST is an acronym for Basic Local Alignment Search Tool, and, as is well
10 known in the art, consists of similarity search programs that interrogate sequence databases for both protein and DNA using heuristic algorithms to seek local alignments. For example, user 101 may conduct a BLAST search using the "blastn" nucleotide sequence database. Results of
15 this batch BLAST search, represented by similar nucleotide and/or protein sequence data 326, may not be available to user 101 for many hours. User 101 may then initiate comparisons and evaluations 332, which may be conducted manually or using various software tools. User
20 101 may subsequently issue report 334 interpreting the findings of the searches and positing strategies and requirements for follow-on experiments.

Inputs to Genomic Portal 400 from User 101

Figure 4 is a functional block diagram showing an
25 illustrative configuration by which user 101 may connect with genomic web portal 400. It will be understood that Figure 4 is simplified and is illustratively only, and that many implementations and variations of the network and Internet connections shown in Figure 4 will be
30 evident to those of ordinary skill in the relevant art.

User 101 employs user computer 100 and analysis applications 199 as noted above, including generating and/or accessing some or all of files 212-217. As shown in Figure 4, files 212-217 are maintained in this example
35 on user database server 412 to which user computer 100 is

- 33 -

implementations, a SIF is a portion of a consensus sequence that has been deemed to best represent the mRNA transcript from a given gene or EST. The consensus sequence may have been derived by comparing and
5 clustering EST's, and possibly also by comparing the EST's to genomic sequence information. A SIF is a portion of the consensus sequence for which probes on the array are specifically designed. With respect to the operations of web portal 400, it is assumed that some
10 microarray probe sets may be designed to detect the expression of genes based upon sequences of EST's.

As was described above, the term "probe set" generally refers to one or more probes from an array of probes on a microarray. For example, in an Affymetrix®
15 GeneChip® probe array, in which probes are synthesized on a substrate, a probe set may consist of 30 or 40 probes, half of which typically are controls. These probes collectively, or in various combinations of some or all of them, are deemed to be indicative of the expression of
20 a gene or EST. In a spotted probe array, one or more spots may similarly constitute a "probe set."

The term "probe-set identifiers" is used broadly herein in that a number of types of such identifiers are possible and are intended to be included within the
25 meaning of this term. One type of probe-set identifier is a name, number, or other symbol that is assigned for the purpose of identifying a probe set. This name, number, or symbol may be arbitrarily assigned to the probe set by, for example, the manufacturer of the probe
30 array. A user may select this type of probe-set identifier by, for example, highlighting or typing the name. Another type of probe-set identifier as intended herein is a graphical representation of a probe set. For example, dots may be displayed on a scatter plot or other
35 diagram wherein each dot represents a probe set.

unique to the consensus sequence or the gene or EST's the consensus sequence is intended to represent. In that case, the sequence specified by the user is a probe-set identifier that identifies the probe set corresponding to the SIF, even though the user-specified sequence is not included in the SIF. Parallel cases are possible with respect to user specifications of partial sequences of EST's and genes or EST's, as those skilled in the relevant art will now appreciate.

10 A further example of a probe-set identifier is an accession number of a gene or EST. Gene and EST accession numbers are publicly available. A probe set may therefore be identified by the accession number or numbers of one or more EST's and/or genes corresponding to the probe set. The correspondence between a probe set and EST's or genes may be maintained in a suitable database, such as that accessed by database application 230 or local library databases 516, from which the correspondence may be provided to the user. Similarly, 20 gene fragments or sequences other than EST's may be mapped (e.g., by reference to a suitable database) to corresponding genes or EST's for the purpose of using their publicly available accession numbers as probe-set identifiers. For example, a user may be interested in 25 product or genomic information related to a particular SIF that is derived from EST-1 and EST-2. The user may be provided with the correspondence between that SIF (or part or all of the sequence of the SIF) and EST-1 or EST-2, or both. To obtain product or genomic data related to 30 the SIF, or a partial sequence of it, the user may select the accession numbers of EST-1, EST-2, or both.

Genomic Web Portal 400

Genomic web portal 400 provides to user 101 data related to one or more genes or EST's. Each gene or EST 35 has at least one corresponding probe set that is

- 37 -

either networked or otherwise. They may be co-located or they may be physically separated. Various operating systems may be employed on any of the computer platforms, possibly depending on the type and/or make of computer platform chosen. Appropriate operating systems include Windows NT®, Sun Solaris, Linux, OS/400, Compaq Tru64 Unix, SGI IRIX, Siemens Reliant Unix, and others.

There may be significant advantages to carrying out the functions of portal 400 on multiple computer platforms in this manner, such as lower costs of deployment, database switching, or changes to enterprise applications, and/or more effective firewalls. Other configurations, however, are possible. For example, as is well known to those of ordinary skill in the relevant art, so-called two-tier or N-tier architectures are possible rather than the three-tier server-side component architecture represented by Figure 5. See, for example, E. Roman, Mastering Enterprise JavaBeans™ and the Java™2 Platform (John Wiley & Sons, Inc., NY, 1999) and J. Schneider and R. Arora, Using Enterprise Java™ (Que Corporation, Indianapolis, 1997), both of which are hereby incorporated by reference in their entireties for all purposes.

It will be understood that many hardware and associated software or firmware components that may be implemented in a server-side architecture for Internet commerce are not shown in Figure 5. Components to implement one or more firewalls to protect data and applications, uninterruptable power supplies, LAN switches, web-server routing software, and many other components are not shown. Similarly, a variety of computer components customarily included in server-class computing platforms, as well as other types of computers, will be understood to be included but are not shown. These components include, for example, processors, memory

- 39 -

using distributed object applications written in the Java™ language. Thus, in one implementation, the functional elements of portal 400 may be written in Java and implemented using J2EE and Enterprise JavaBeans™.

5 Various other software development approaches or architectures may be used to implement the functional elements of portal 400 and their interconnection, as will be appreciated by those of ordinary skill in the art.

One implementation of these platforms and components
10 is shown in Figure 6. Figure 6 is a simplified graphical representation of illustrative interactions between user-side internet client 410 on the user side and input and output managers 532 and 534 of Internet server 530 on the portal side, as well as communications among the three
15 tiers (servers 510, 520, and 530) of portal 400. Browser 605 on client 410 sends and receives HTML documents 620 to and from server 530. HTML document 625 includes applet 627. Browser 605, running on user computer 103, provides a run-time container for applet 627. Functions of
20 managers 532 and 534 on server 530, such as the performance of GUI operations, may be implemented by servlet and/or JSP 640 operating with a Java™ platform. A servlet engine executing on server 530 provides a runtime container for servlet 640. JSP (Java Server
25 Pages) from Sun Microsystems, Inc. is a script-like environment for GUI operations; an alternative is ASP (Active Server Pages) from the Microsoft Corporation. App server 650 is the middleware product referred to above, and executes on application server 520. EJB
30 (Enterprise JavaBeans™ is a standard that defines an architecture for enterprise beans, which are application components. CORBA (Common Object Request Broker Architecture) similarly is a standard for distributed object systems, i.e., the CORBA standards are implemented
35 by CORBA-compliant products such as Java™ IDL. An

- 41 -

coordinates the storage, maintenance, supplementation, and all other transactions from or to any of local databases 511, 513, 514, 516, and 518. Manager 512 may undertake these functions in cooperation with appropriate 5 database applications such as the Oracle® 8.0.5 database management system.

In some implementations, manager 512 periodically updates local genomic database 518. The data updated in database 518 includes data related to genes or EST's that 10 correspond with one or more probe sets. The probe sets may be those used or designed for use on any microarray product, and/or that are expected or calculated to be used in microarray products of any manufacturer or researcher. For example, the probe sets may include all 15 probe sets synthesized on the line of stocked GeneChip® probe arrays from Affymetrix, Inc., including its Arabidopsis Genome Array, CYP450 Array, Drosophila Genome Array, E. coli Genome Array, GenFlex™ Tag Array, HIV PRT Plus Array, HuGeneFL Array, Human Genome U95 Set, HuSNP 20 Probe Array, Murine Genome U74 Set, P53 Probe Array, Rat Genome U34 Set, Rat Neurobiology U34 Set, Rat Toxicology U34 Array, or Yeast Genome S98 Array. The probe sets may also include those synthesized on custom arrays for user 101 or others. However, the data updated in database 518 25 need not be so limited. Rather, it may relate to any number of genes or EST's. Types of data that may be stored in database 518 are described below in relation to the operations of manager 522 in directing the periodic collection of this data from remote sources providing the 30 locally maintained data in database 518 to users.

Database 516 includes data of a type referred to above in relation to database application 230, i.e., data that associates probe sets with their corresponding gene or EST and their identifiers. Database 516 may also 35 include SIF's, and other library data. User-service

current orders, past orders, and so on, may be obtained from users, all as will be readily apparent to those of ordinary skill in the art. Also, information related to users may be developed by recording and/or analyzing the interactions of users with portal 400, in accordance with known techniques used in e-commerce. For example, user-service manager 522 may take note of users' areas of genomic interest, their purchase or product-inquiry activities, the frequency of their accessing of various services, and so on, and provide this information to database manager 512 for storage or update in database 519.

Another functional element of portal 400 is user-service manager 522. Manager 522 may periodically cause database manager 512 to update local genomic database 518 from various sources, such as remote databases 402. For example, according to any chronological schedule (e.g., daily, weekly, etc.), manager 522 may, in accordance with known techniques, initiate searches of remote databases 402 by formulating appropriate queries, addressed to the URL's of the various databases 402, or by other conventional techniques for conducting data searches and/or retrieving data or documents over the Internet. These search queries and corresponding addresses may be provided in a known manner to output manager 534 for presentation to databases 402. Input manager 532 receives replies to the queries and provides them to manager 522, which then provides them to database manager 512 for updating of database 518, all in accordance with any of a variety of known techniques for managing information flow to, from, and within an Internet site.

Portal application manager 526 manages the administrative aspects of portal 400, possibly with the assistance of a middleware product such as an applications server product. One of these administrative

- 45 -

data, marker structure or location data, polymorphism data, homology data, protein-family classification data, pathway data, alternative-gene naming data, literature-recitation data, and annotation data. Many other
5 examples are possible. Also, genomic data not currently available but that becomes available in the future may be accessed and locally maintained as described herein. Examples of remote databases 402 currently suitable for accessing in the manner described include GenBank,
10 GenBank New, SwissProt, GenPept, DB EST, Unigene, PIR, Prosite, PFAM, Prodom, Blocks, PDB, PDBfinder, EC Enzyme, Kegg Pathway, Kegg Ligand, OMIM, OMIM Map, OMIM Allele, DB SNP, and PubMed. Hundreds of other databases currently exist that are suitable, and thus this list is
15 merely illustrative.

Moreover, local genomic database 518 may also be supplemented with data obtained or deduced (by user-service manager 522) from other of the local databases serviced by database manager 512. In particular,
20 although local products database 514 is shown for convenience of illustration as separate from database 518, it may be the same database. Alternatively, or all or part of the data in database 514 may be duplicated in, or accessible from, database 518.

25 More specific examples are now provided of how user service manager 522 may receive and respond to requests from user 101 for genomic information and for product information and/or ordering. These examples are described in relation to Figures 7, 8 and 9.

30 Figure 7 is a flow chart representing an illustrative method by which the illustrated embodiment of portal 400 may respond to a user's request for genomic or product information. In accordance with step 710 of this example, input manager 532 receives from client 410
35 over Internet 499 a request by user 101 for data. This

- 47 -

statistics or information regarding user 101's request may be recorded, as noted above.

In accordance with step 725, user-service manager 522 formulates an appropriate query (using, for example, a version of the SQL language) for correlating probe-set identifiers with corresponding genes or EST's. Gene or EST determiner 820 is the functional element of manager 522 that illustratively executes this task. Determiner 820 forward the query to database manager 512. If the probe-set identifiers provided by user 101 include sequence information, then the query may seek from database 511, and/or from SIF information in database 516, the identity of the one or more probe sets having a corresponding (e.g., similar in biological significance) sequence. If the probe-set identifiers include names or numbers (e.g., accession numbers), then the query may seek the identity of the probe sets from database 516 that, as noted, includes data that associates names, numbers, and other probe-set identifiers with corresponding genes or EST's. User 101 may also have locally employed database application 230 to obtain this information, and included it in the information request in accordance with known techniques. In this case, step 725 need not be performed.

As indicated in step 730, user-service manager 522 may then correlate the indicated genes and/or EST's with genomic information and/or product information. The performance of this task is undertaken by correlator 830 in the illustrated example. In one of many possible implementations, correlator 830 formulates a query via database manager 512 to database 513 in order to obtain links to appropriate information in local products database 514 and/or local genomic database 518. Figure 9 is a simplified graphical representation of database 513. Those of ordinary skill in the art will appreciate that

- 49 -

Following the appropriate links 904 to probe-set ID's 912, one or more links 916 to related products and/or genomic data may be obtained. For example, link 904N may link to probe-set 912C, which is associated with 5 links 916C to related product and/or genomic data. The information used to establish this association may be predetermined based on expert input and/or computer-implemented analysis (e.g., statistical and/or by an adaptive system such as a neural network) of the nature 10 of inquiries by users. For example, it may be observed or anticipated (by humans or computers, as noted) that users conducting gene expression experiments resulting in the identification of certain genes may wish to use antibodies against the genes to conduct follow-on protein 15 level experiments. The association between the genes and the appropriate antibodies may be stored in an appropriate database, such as database 516. Links 916C may thus include links to product or genomic data identifiers that identify links to data about the 20 appropriate antibodies (for example, a link to product/genomic ID 922A), to catalogues of antibodies generally (e.g., ID 922B), or to a probe array specifically designed for detecting alternatively spliced forms of the genes of interest (e.g., ID 922C). It is 25 assumed for illustrative purposes that, in a particular aspect of this example, link 916C leads to ID 922C. Information about the availability of splice-variant probe arrays may be predetermined by the contents of links 926. For example, links 926D (associated with ID 30 922C, as shown) may be stored Internet and/or database-query URL's leading to vendor web pages, local products database 514, and/or local genomic database 518. Also, the content of links 926D may be dynamically determined by query of databases 514 or 518 or of remote data 35 sources such as databases 402 or web pages 404. These

- 51 -

provided to output manager 534 for transmission to client 410.

In some implementations, user 101 may respond to the data thus transmitted by indicating a desire to purchase a product or receive further information. A request for further information may be processed in a manner similar to that described above with respect to Figure 7. If user 101 indicates a desire to purchase a product (see decision element 745), the indicated product may be prepared for shipment or otherwise processed, and the user's account may be adjusted, in accordance with known techniques for conducting e-commerce. As one of many alternative implementations, user-service manager 522 may notify the product vendor of user 101's order and the vendor may ship, or order the shipment of, the product. Manager 522 may then note, in one aspect of this implementation, that a fee should be collected from the vendor for the referral.

In some implementations of portal 400, user 101 may provide to portal 400 (e.g., via client 410, Internet 499, and input manager 532) one or more gene or EST ascension numbers or other gene or EST identifiers. Alternatively, or in addition, user 101 may provide to portal 400 one or more probe-set identifiers. User 101 may obtain the gene, EST, and/or probe-set identifier from a public source, from notations user 101 has taken as a result of experiments with a probe array or otherwise, from a list of genes or EST's having corresponding probes on a probe array, or from any other source or obtained in any other manner. Input manager 532 receives the one or more gene, EST, or probe-set identifiers and provides it or them to user-service manager 522, which formulates a query to database manager 512. In accordance with known query techniques and formats, the query seeks information from local products

- 53 -

with one or more product data and that causes (e.g., via database manager 512) the product data to be obtained either locally from, e.g., database 514 or, in some implementations, remotely from, e.g., pages 404 or 5 databases 402; and output manager 534 that provides the product data to user 101.

Similarly, a method is provided for providing biological product data, including the steps of: receiving from user 101 any one or more of a gene, EST, 10 and/or probe-set identifier; correlating the gene, EST, and/or probe-set identifier with one or more product data; causing the product data to be obtained either locally from, e.g., database 514 and/or remotely from, e.g., pages 404 or databases 402; and providing the 15 product data to user 101.

As indicated above, functional elements of portal 400 may be implemented in hardware, software, firmware, or any combination thereof. In the embodiment described above, it generally has been assumed for convenience that 20 the functions of portal 400 are implemented in software.

That is, the functional elements of the illustrated embodiment comprise sets of software instructions that cause the described functions to be performed. These software instructions may be programmed in any 25 programming language, such as Java, Perl, C++, another high-level programming language, low-level languages, and any combination thereof. The functional elements of portal 400 may therefore be referred to as carrying out "a set of genomic web portal instructions," and its 30 functional elements may similarly be described as sets of genomic web portal instructions for execution by servers 510, 520, and 530.

In some embodiments, a computer program product is described comprising a computer usable medium having 35 control logic (computer software program, including

- 55 -

For example, the functions of processors 842 and 844 could be ascribed to a single functional element. Similarly, some or all of the functions of database manager 512 could be carried out by user-service manager 5 522, and/or by input manager 532.

Also, the sequencing of functions or portions of functions generally may be altered. For example, the functions of account ID determiner 810 may be carried out after those of user data processor 840. The flow of data 10 and control in Figure 8 in this regard thus is exemplary only. Similarly, the method steps shown in Figure 7 need not always be carried out in the order suggested by the illustrative example of that figure. For instance, method step 720 of identifying the user could be carried 15 out after that of steps 725, 730, or 735.

Certain functional elements, files, data structures, and so on, may be described in the illustrated embodiments as located in system memory 120 of computer 100 or generally in servers 510, 520, or 530. In other 20 embodiments, however, they may be located on, or distributed across, computer systems or other platforms that are co-located and/or remote from each other. For example, any one or more of data files or data structures 511, 513, 514, 516, or 518, shown in Figure 5 as co- 25 located on and "local" to server 510, may be located in a computer system or systems remote from server 510. In those cases, the operations of database manager 512 with respect to these data files or data structures may be carried out over a network or by any of numerous other 30 known means for transferring data and/or control to or from a remote location.

In addition, it will be understood by those skilled in the relevant art that control and data flows between and among functional elements and various data structures 35 may vary in many ways from the control and data flows

- 57 -

CLAIMS

1. A system for providing data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of a biological molecule, comprising:
 - an input manager constructed and arranged to receive from a user a selection of a first set of one or more of the probe-set identifiers;
 - 10 a gene determiner constructed and arranged to identify a first set of one or more genes or EST's corresponding to the probe sets identified by the first set of probe-set identifiers;
 - a correlator constructed and arranged to correlate 15 the first set of genes or EST's with a first set of one or more data; and
 - an output manager constructed and arranged to provide the first set of data to the user.
- 20 2. The system of claim 1, wherein:
 - the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of nucleic acid.
- 25 3. The system of claim 1, wherein:
 - the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of mRNA transcripts of corresponding genes.
- 30 4. The system of claim 1, wherein:
 - the first set of probe-set identifiers comprises all or part of a second set of one or more probe-set identifiers of probe sets that have enabled detection of

13. The system of claim 1, wherein:
at least a first probe-set identifier of the first
set of probe-set identifiers comprises a gene identifier
5 of the gene corresponding to the first probe-set
identifier.
14. The system of claim 13, wherein:
the gene identifier comprises an accession number.
- 10 15. The system of claim 1, wherein:
the user selects the first set of probe-set
identifiers based, at least in part, on an indication of
a degree of expression or differential expression of the
15 genes or EST's corresponding to the probe sets identified
by the first set of probe-set identifiers.
16. The system of claim 1, wherein:
the first set of one or more data includes one or
20 any combination of product data related to availability,
pricing, composition, suitability, or ordering.
17. The system of claim 16, wherein:
the first set of one or more data includes product
25 data regarding a biological device or substance, or a
reagent that may be used with a biological device or
substance.
18. The system of claim 17, wherein:
30 the device, substance, or reagent includes one or
any combination of an oligonucleotide, probe array,
clone, antibody, or protein.
19. The system of claim 1, wherein:

26. The method of claim 25, wherein:
the first set of probe-set identifiers identify
probe sets that are capable of enabling the detection of
5 a biological molecule that consists of nucleic acid.

27. The method of claim 25, wherein:
the first set of probe-set identifiers identify
probe sets that are capable of enabling the detection of
10 a biological molecule that consists of mRNA transcripts
of corresponding genes.

28. A computer program product for providing data
related to one or more genes or EST's, wherein each gene
15 or EST has at least one corresponding probe set
identified by a probe-set identifier and capable of
enabling detection of a biological molecule, wherein the
computer program product, when executed on a computer
system, performs a method comprising the steps of:
20 receiving from a user a selection of a first set of
one or more of the probe-set identifiers;
identifying a first set of one or more genes or
EST's corresponding to the probe sets identified by the
first set of probe-set identifiers;
25 correlating the first set of genes or EST's with a
first set of one or more data; and
providing the first set of data to the user.

29. The computer program product of claim 28, wherein:
30 the first set of probe-set identifiers identify
probe sets that are capable of enabling the detection of
a biological molecule that consists of nucleic acid.

30. The computer program product of claim 28, wherein:

the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of mRNA transcripts of corresponding genes.

5

34. The system of claim 31, wherein:

at least one of the probe sets identified by the first set of probe-set identifiers is disposed on a GeneChip® probe array.

10

35. A system for providing data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of a

15 biological molecule, comprising:

an input manager constructed and arranged to receive from a user a selection of a first set of one or more of the probe-set identifiers;

20

a gene determiner constructed and arranged to identify a first set of one or more genes or EST's corresponding to the probe sets identified by the first set of probe-set identifiers;

25

an account identification determiner constructed and arranged to identify an account corresponding to the user;

a correlator constructed and arranged to correlate the first set of genes or EST's with a first set of one or more product data including product pricing data;

30

an account data processor constructed and arranged to adjust the account corresponding to the user based, at least in part, on the product pricing data; and

an output manager constructed and arranged to provide the first set of product data to the user.

an output manager constructed and arranged to provide at least a portion of the first set of product data to the user.

5 40. The system of claim 39, wherein:

the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of nucleic acid.

10 41. The system of claim 39, wherein:

the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of mRNA transcripts of corresponding genes.

15

42. The system of claim 39, wherein:

the input manager further is constructed and arranged to receive from the user a second user selection of one or more products for purchase based on the first
20 set of product data.

43. The system of claim 42, further comprising:

an account data processor constructed and arranged to adjust the account corresponding to the user based, at
25 least in part, on the product pricing data corresponding to the second user selection.

- 67 -

identifying an account corresponding to the user;
and

adjusting the account corresponding to the user
based, at least in part, on the product pricing data
5 corresponding to the second user selection.

49. A method for placing a computer-implemented
inquiry or order regarding purchase of one or more
products, comprising the steps of:

10 receiving at a user computer a first user selection
of a first set of one or more probe-set identifiers,
wherein each probe-set identifier identifies a probe set
that has enabled detection of a biological molecule;
providing the first user selection over the Internet
15 to a portal system capable of correlating product data
with one or more genes or EST's corresponding to the
probe sets identified by the first set of probe-set
identifiers; and
receiving the correlated product data from the
20 portal system.

50. The method of claim 49, wherein:
the first set of probe-set identifiers identify
probe sets that are capable of enabling the detection of
25 a biological molecule that consists of nucleic acid.

51. The method of claim 49, wherein:
the first set of probe-set identifiers identify
probe sets that are capable of enabling the detection of
30 a biological molecule that consists of mRNA transcripts
of corresponding genes.

56. The system of claim 53, wherein:
the database manager updates the local genomic
database according to a chronological period.
- 5
57. The system of claim 56, wherein:
the chronological period is predetermined.
58. The system of claim 56, wherein:
10 the chronological period is greater than about ten
hours and less than about ten days.
59. The system of claim 53, wherein:
the database manager periodically updates the local
15 genomic database with update data consisting of any
combination of one or more of sequence data, exonic
structure or location data, splice-variants data, marker
structure or location data, polymorphism data, homology
data, protein-family classification data, pathway data,
20 alternative-gene naming data, literature-recitation data,
or annotation data.
60. The system of claim 53, wherein:
the database manager periodically updates the local
25 genomic database with update data from one or more remote
databases.
61. The system of claim 60, wherein:
the updating from one or more remote databases
30 comprises updating over the Internet.
62. The system of claim 61, wherein:
the remote databases consist of any combination of
one or more of GenBank, GenBank New, SwissProt, GenPept,
35 DB EST, Unigene, PIR, Prosite, PFAM, Prodom, Blocks, PDB,

- 71 -

the input manager receives the remote user's selection over a network.

70. The system of claim 69, wherein:
5 the network includes the Internet.

71. The system of claim 53, wherein:
the user includes a remote user, and
the output manager provides the first set of data to
10 the user over a network.

72. The system of claim 71, wherein:
the network includes the Internet.

15 73. The system of claim 53, wherein:
at least one of the probe-set identifiers comprises
a gene identifier of the gene corresponding to the probe-
set identifier.

20 74. The system of claim 73, wherein:
the gene identifier comprises an accession number.

75. A system for providing data related to one or
more genes or EST's, wherein each gene or EST has a
25 corresponding probe set identified by a probe-set
identifier and capable of enabling detection of the
expression of the gene, the system comprising:
a database manager constructed and arranged to
periodically update a local genomic database comprising
30 data related to the genes or EST's, wherein the updating
is done according to a predetermined period;
an input manager constructed and arranged to
dynamically receive a user-initiated selection of a first
set of one or more of the probe-set identifiers;

- 73 -

to the genes or EST's with update data from one or more remote databases, wherein the updating is done over the Internet according to a predetermined period;

an input manager constructed and arranged to
5 dynamically receive a user-initiated selection of a first set of one or more of the probe-set identifiers;

a user-service manager constructed and arranged to construct from the local genomic database a first set of data related to genes or EST's corresponding to the first
10 set of probe-set identifiers;

an output manager constructed and arranged to provide the first set of data to the user.

78. A system for providing data related to one or
15 more genes or EST's, wherein each gene or EST has a corresponding probe set identified by a probe-set identifier and capable of enabling detection of the expression of the gene, the system comprising:

a database manager constructed and arranged to
20 update a local genomic database comprising data related to the genes or EST's with update data from one or more remote databases, wherein the updating is done over the Internet according to a predetermined period;

an input manager constructed and arranged to
25 dynamically receive over the Internet a user-initiated selection of a first set of one or more of the probe-set identifiers;

a user-service manager constructed and arranged to construct from the local genomic database a first set of
30 data related to genes or EST's corresponding to the first set of probe-set identifiers; and

an output manager constructed and arranged to provide over the Internet the first set of data to the user.

- 75 -

periodically updating a local genomic database comprising data related to the genes or EST's;

receiving from a user a selection of a first set of one or more of the probe-set identifiers;

- 5 constructing from the local genomic database a first set of data related to genes or EST's corresponding to the first set of probe-set identifiers; and
providing the first set of data to the user.

10 82. A system for providing product data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of a biological molecule, comprising:

- 15 an input manager constructed and arranged to receive from a user a selection of a first set of one or more of the probe-set identifiers;

a correlator constructed and arranged to correlate the first set of probe-set identifiers with a first set

- 20 of one or more product data; and

an output manager constructed and arranged to provide the first set of data to the user.

83. The system of claim 82, wherein:

- 25 the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of nucleic acid.

84. The system of claim 84, wherein:

- 30 the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of a biological molecule that consists of mRNA transcripts of corresponding genes.

- 35 85. The system of claim 84, wherein:

- 77 -

92. The system of claim 82, wherein:

the first set of one or more product data includes data stored, at least in part, in a local products database.

5

93. The system of claim 82, wherein:

the first set of one or more data includes at least one link to remote data representing a vendor of biological products.

10

94. A method for providing product data related to one or more genes or EST's, wherein each gene or EST has at least one corresponding probe set identified by a probe-set identifier and capable of enabling detection of
15 a biological molecule, comprising the steps of:

receiving from a user a selection of a first set of one or more of the probe-set identifiers;

correlating the first set of probe-set identifiers with a first set of one or more product data; and

20 providing the first set of data to the user.

95. The method of claim 94, wherein:

the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of
25 a biological molecule that consists of nucleic acid.

96. The method of claim 94, wherein:

the first set of probe-set identifiers identify probe sets that are capable of enabling the detection of
30 a biological molecule that consists of mRNA transcripts of corresponding genes.

97. The method of claim 94, wherein:

the probe sets identified by the first set of probe-
35 set identifiers are disposed on one or more probe arrays.

correlating the gene or EST identifiers with one or more product data; and
providing the product data to the user.

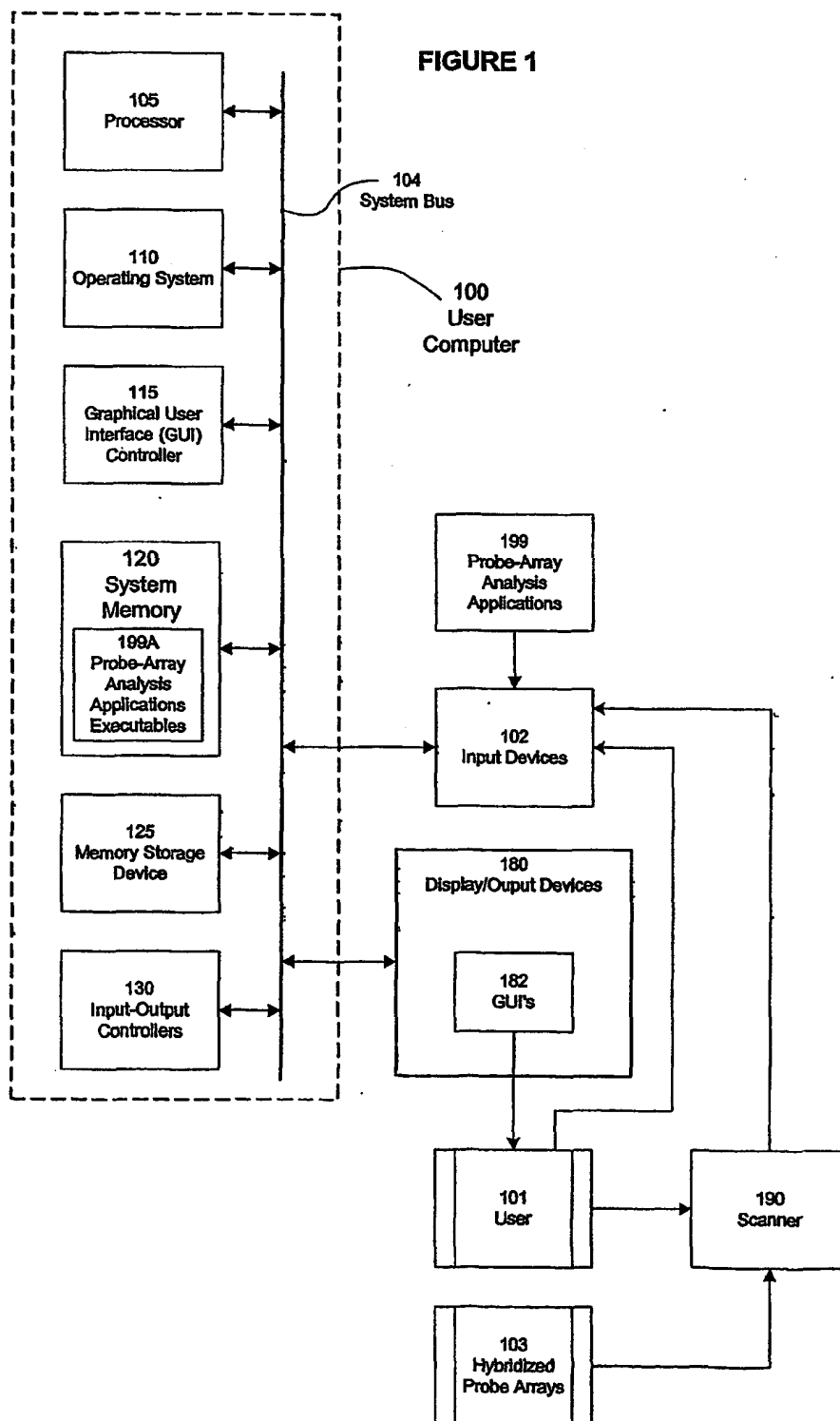
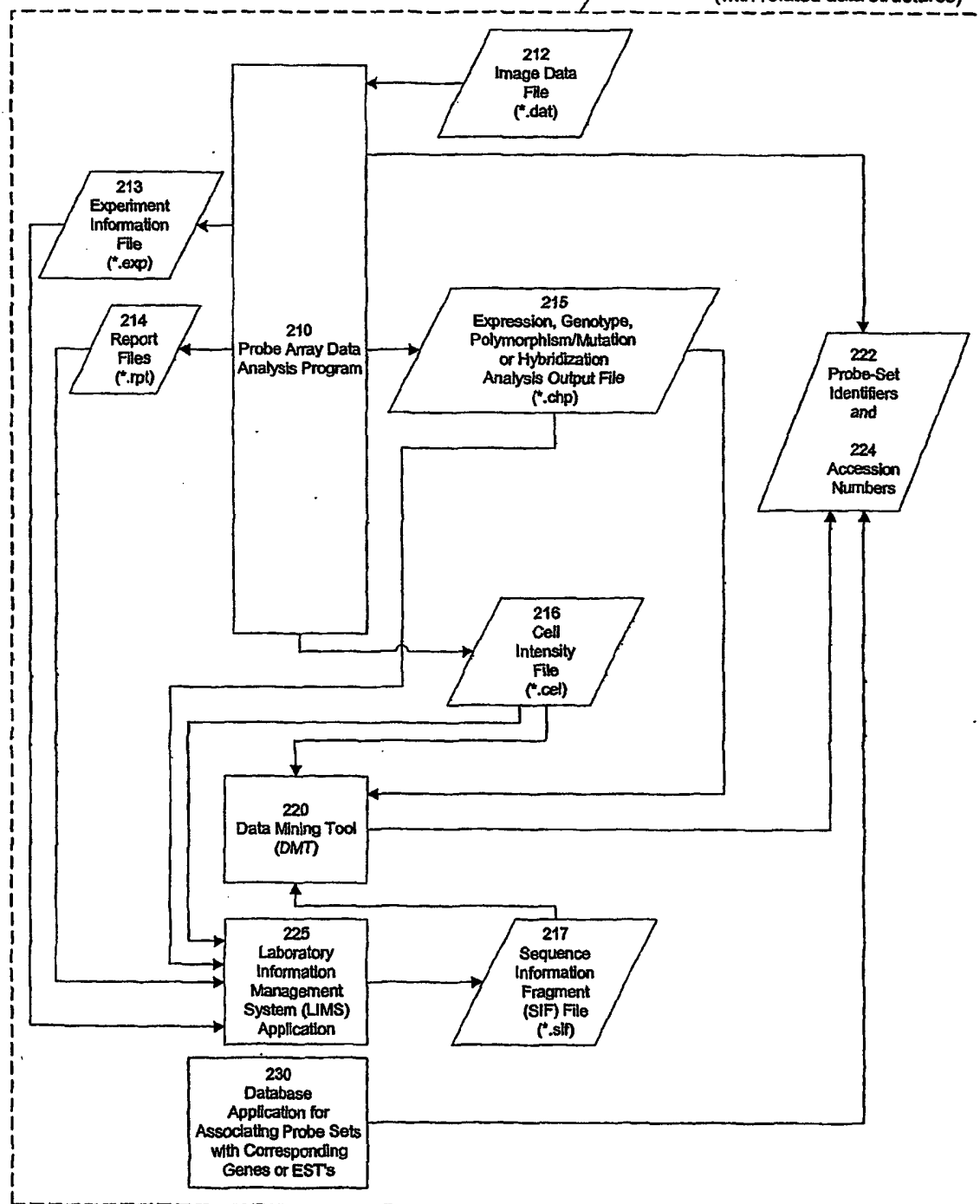


FIGURE 2

199A
Probe-Array Analysis
Applications Executables
(with related data structures)



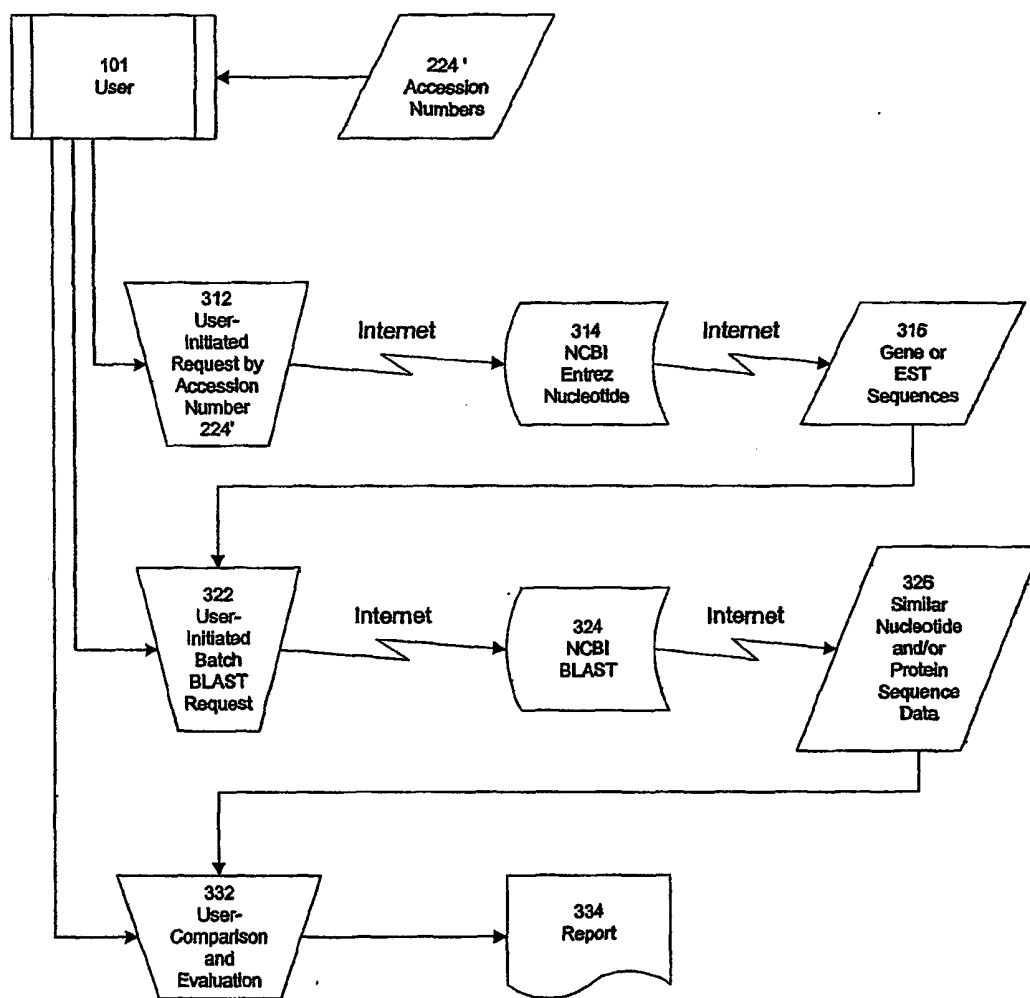
**FIGURE 3
(PRIOR ART)**

FIGURE 4

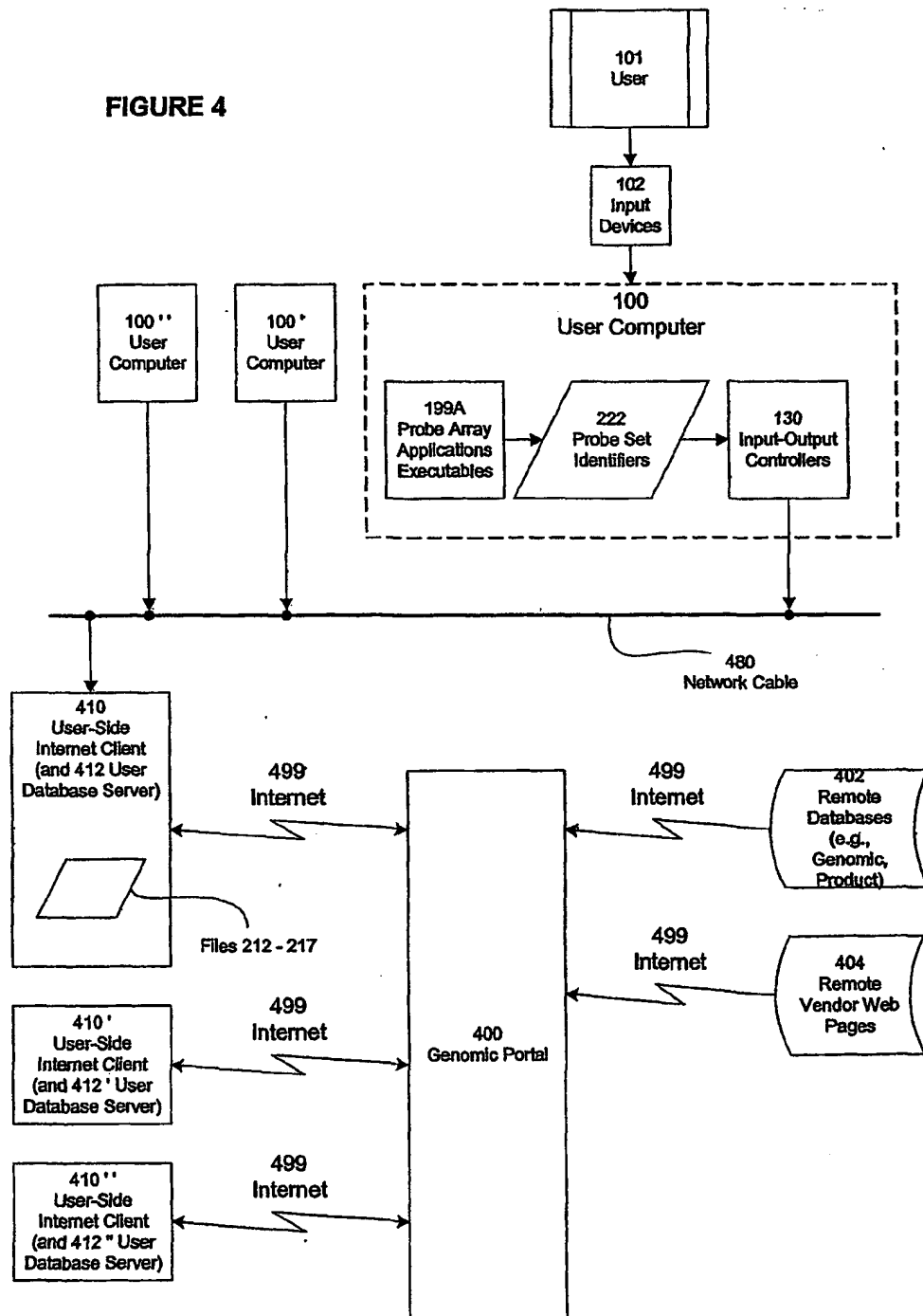


FIGURE 5

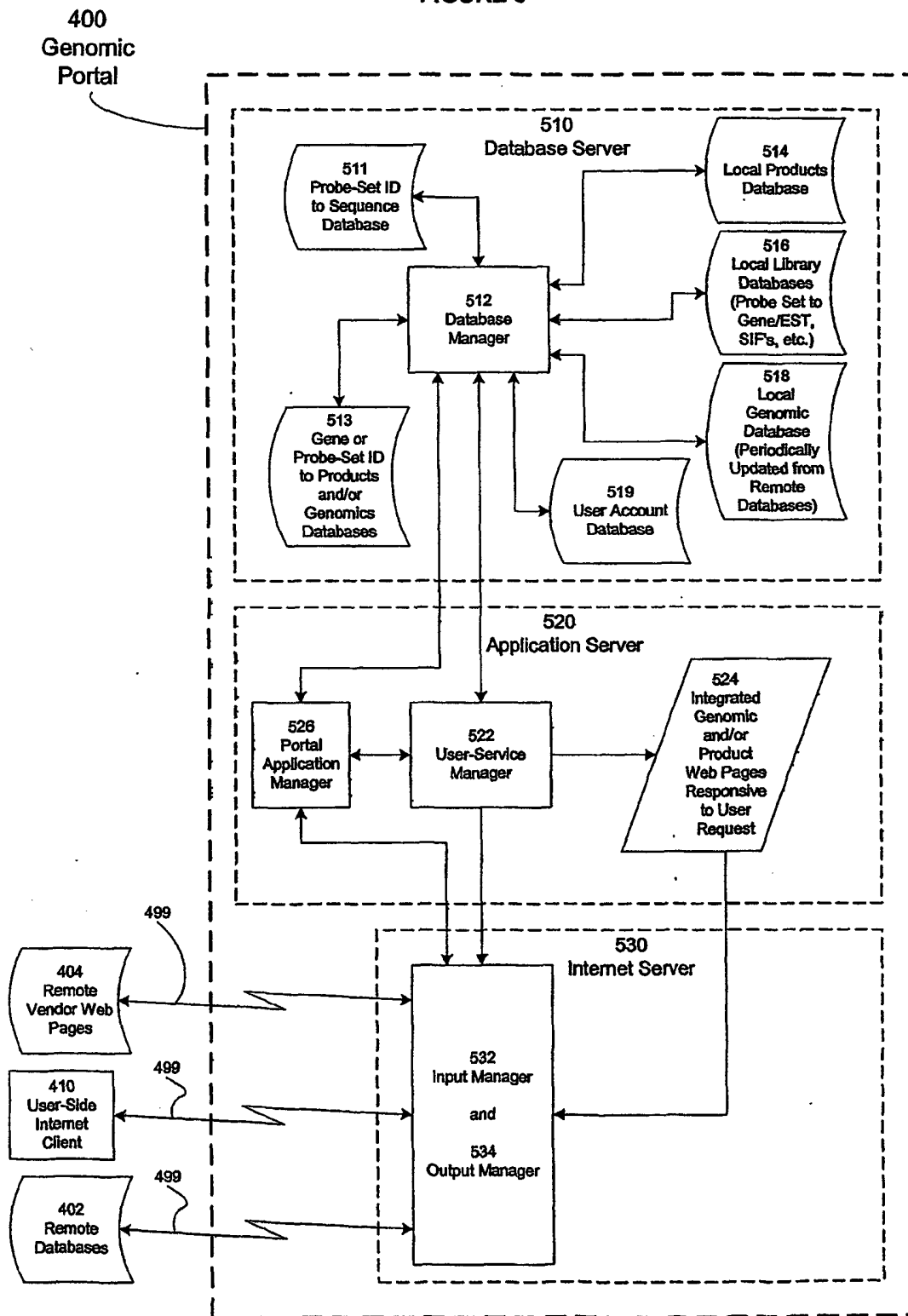


FIGURE 6

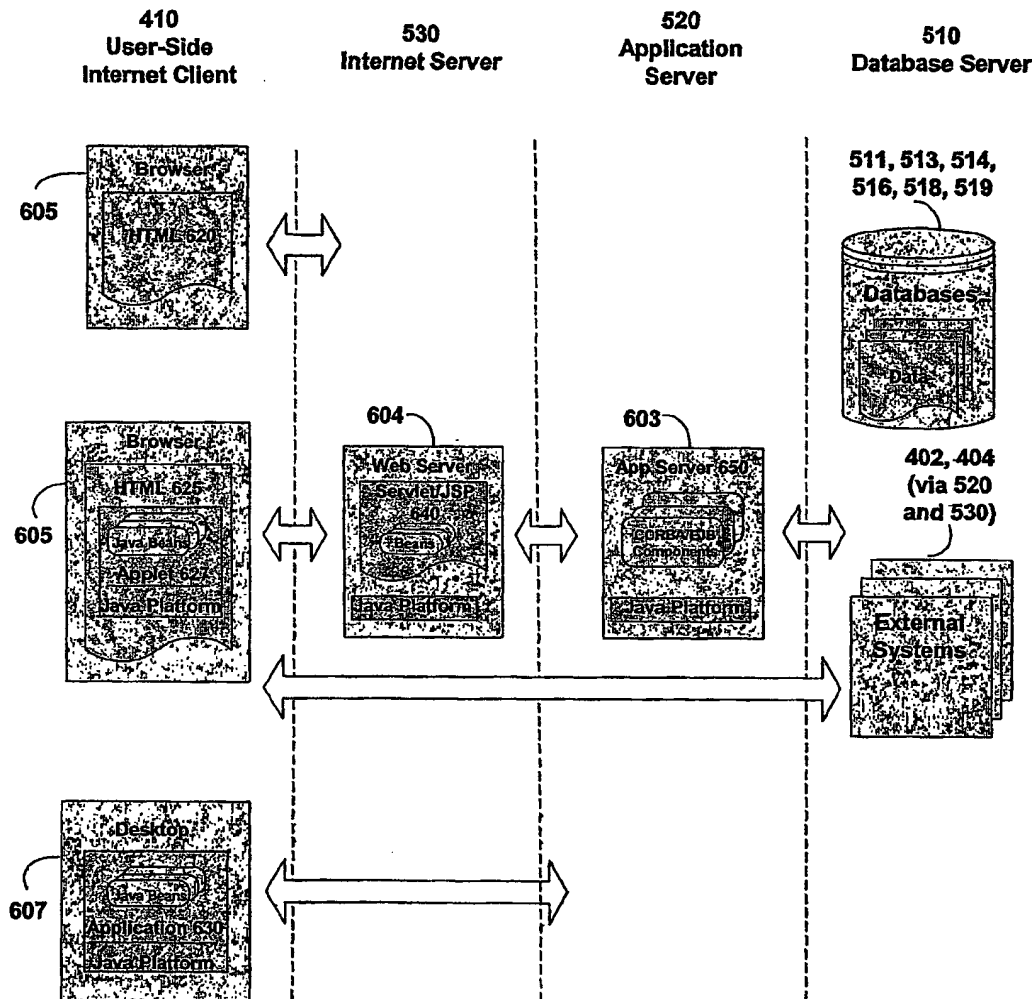


FIGURE 7

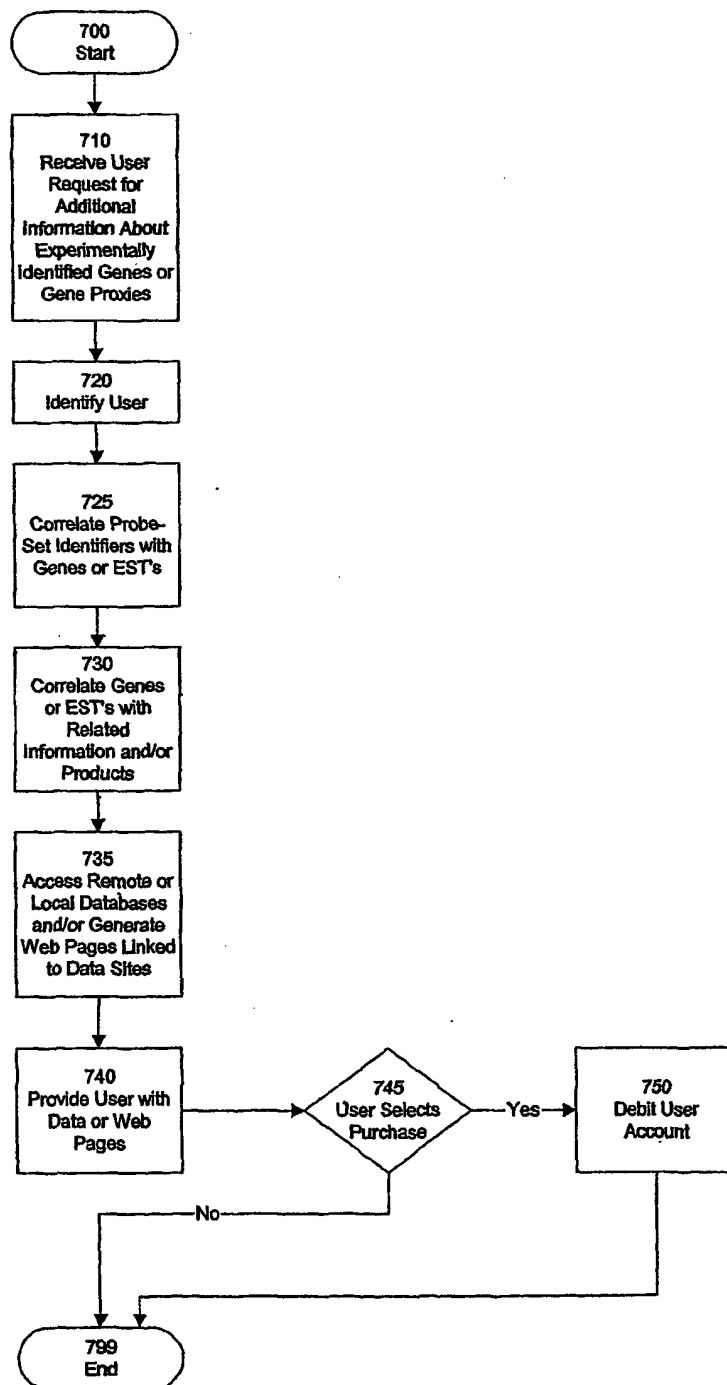


FIGURE 8

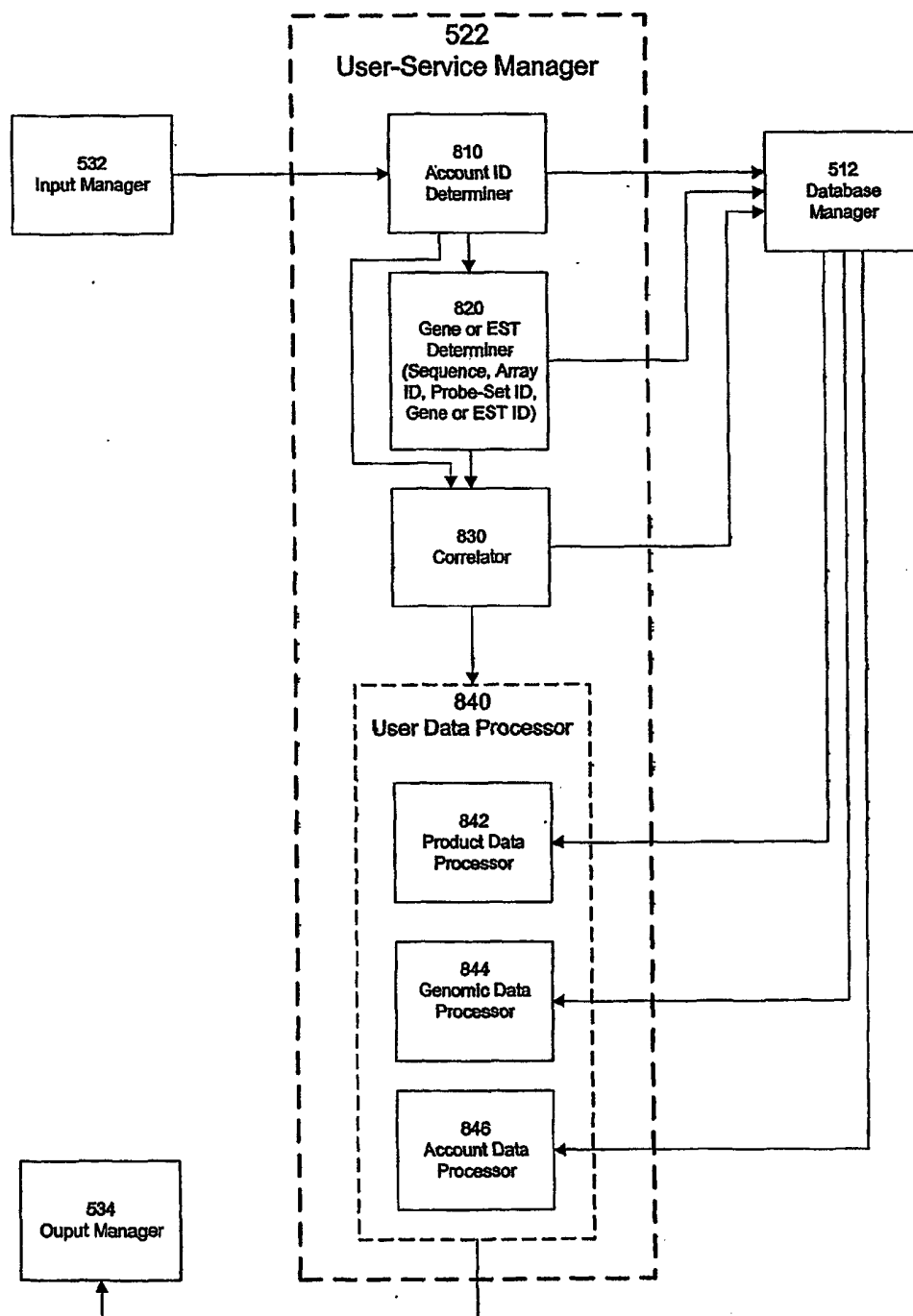


FIGURE 9

513
Gene or Probe-Set ID
to Products and/or
Genomics Databases

902 Gene or EST ID	904 Link to Probe-Set IDS 912
ID 902A	Link 904N
ID 902B	Link 904N
ID 902C	Link 904M

912 Probe-Set ID (accession #, sequence, etc.)	914 Array ID	916 Links to Related Product and/or Genomic Data IDs
Probe-Set 912A	Array 914N	Links 916A
Probe-Set 912B	Array 914N	Links 916B
Probe-Set 912C	Array 914P	Links 916C
Probe-Set 912D	Array 914Q	Links 916D
Probe-Set 912E	Array 914N	Links 916E

922 Product and/or Genomic Data ID	924 Link to Local Products Database 514 and/ or Genomic Database 518	926 Links to Vendor or Genomic Database Web Pages
Product/Genomic ID 922A	Link 924A	Links 926A
Product/Genomic ID 922B	Link 924B	Links 926B
Product/Genomic ID 922C	Link 924C	Links 926C
Product/Genomic ID 922C	Link 924D	Links 926D
Product/Genomic ID 922D	Link 924E	Links 926E

FIGURE 10

1010

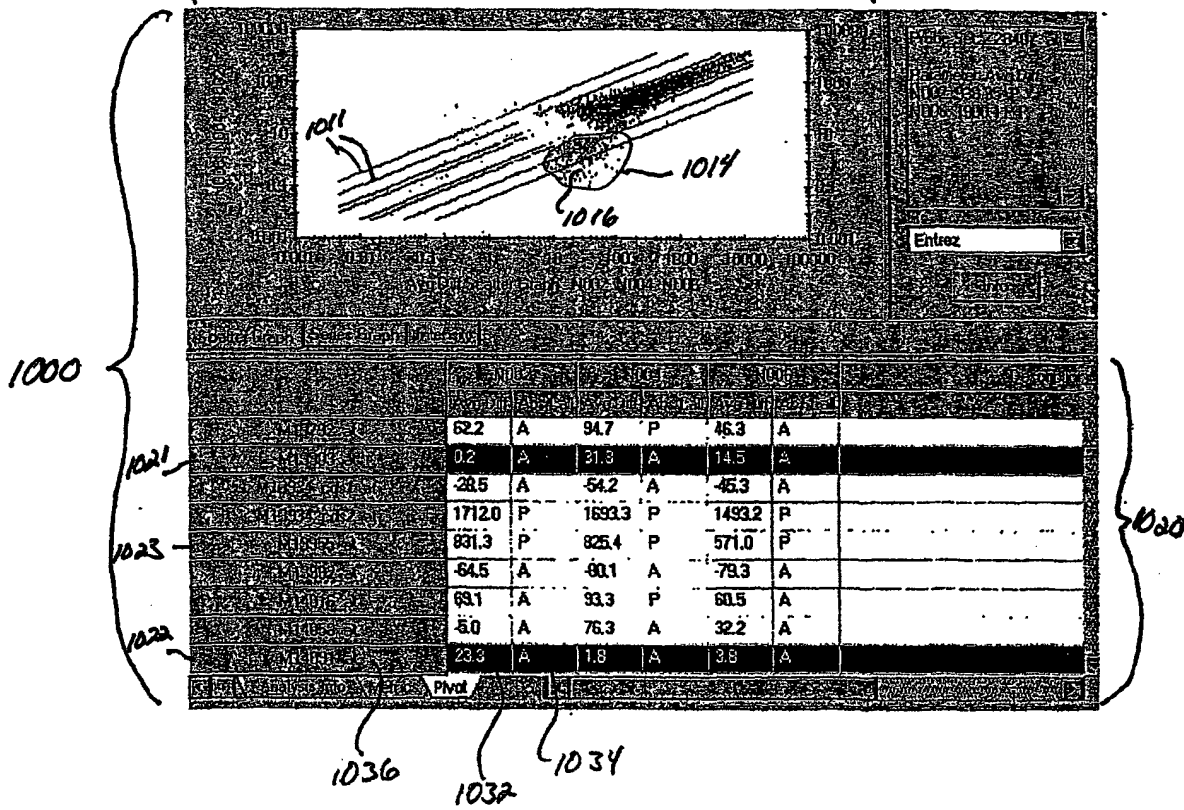


FIGURE 11

